



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 6 :</b> C12Q 1/66, 1/00, C12P 21/06, C07H 21/04, G01N 33/53, 33/567, 33/573, 33/58, 33/60, 33/531, 33/533, 21/00, 21/76, 21/75, 31/00	<b>A1</b>	<b>(11) International Publication Number:</b> WO 00/14271 <b>(43) International Publication Date:</b> 16 March 2000 (16.03.00)
<b>(21) International Application Number:</b> PCT/US99/20207 <b>(22) International Filing Date:</b> 2 September 1999 (02.09.99)  <b>(30) Priority Data:</b> 60/099,068      3 September 1998 (03.09.98)      US 60/135,835      24 May 1999 (24.05.99)      US  <b>(71) Applicant (for all designated States except US):</b> LOMA LINDA UNIVERSITY [US/US]; Loma Linda, CA 92350 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> SZALAY, Aladar, A. [US/US]; 7327 Fairwood Lane, Highland, CA 92346 (US). WANG, Yubao [CN/US]; 24931 Prospect Avenue, Loma Linda, CA 92354 (US). WANG-PRUSKI, Gefu [CA/CA]; 114 Hillcrest Avenue, Truro, Nova Scotia B2N 4L3 (CA).  <b>(74) Agents:</b> SHELDON, Jeffrey, G. et al.; Sheldon & Mak, 9th floor, 225 South Lake Avenue, Pasadena, CA 91101 (US).	<b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i>	
<b>(54) Title:</b> METHOD FOR STUDYING PROTEIN INTERACTIONS <i>IN VIVO</i>  <b>(57) Abstract</b>  A method for determining whether a first protein interacts with a second protein within a living cell. The method comprises providing the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore within the cell. The complexed first protein and the complexed second protein are allowed to come into proximity to each other within the cell. Then, any fluorescence from the acceptor fluorophore resulting from luminescence resonance energy transfer from the donor luciferase is detected, where fluorescence from the acceptor fluorophore indicates that the first protein has interacted with the second protein.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**METHOD FOR STUDYING PROTEIN INTERACTIONS *IN VIVO***

5

10

15

**BACKGROUND**

The study of interactions between proteins in living cells is often necessary to understand the proteins' functions and their mechanisms of action. These interactions are currently studied using immuno-precipitation, the yeast two hybrid method, and  $\beta$ -gal complementation method.

20

However, these methods are associated with several disadvantages. For example, these methods are associated with false positives. Second, they do not permit the determination of quantitative information regarding the interactions. Further, they do not allow for *in vivo* real time monitoring of the interactions.

25

Therefore, it would be advantageous to have another method of studying interactions between proteins *in vivo*, which does not have these disadvantages. Further preferably, the method could be used with a wide variety of proteins and in a wide variety of living cells. Also preferably, the method could be used to determine the interactions between molecules other than proteins.

**SUMMARY**

30

According to one embodiment of the present invention, there is provided a

method for determining whether a first protein interacts with a second protein within a living cell. The method comprises providing the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore within the cell. The donor luciferase is capable of luminescence resonance energy transfer to the acceptor fluorophore when the first protein is in proximity to the second protein. Then, the complexed first protein and the complexed second protein are allowed to come into proximity to each other within the cell. Next, any fluorescence from the acceptor fluorophore is detected. Fluorescence of the acceptor fluorophore resulting from luminescence resonance energy transfer from the donor luciferase to acceptor fluorophore the indicates that the first protein has interacted with the second protein.

In a preferred embodiment, providing the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore comprises genetically engineering DNA and transferring the genetically engineered DNA to the living cell causing the cell to produce the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore. In a particularly preferred embodiment, the cell which is provided with the first protein complexed to a donor luciferase and the cell which is provided with the second protein complexed to an acceptor fluorophore are mammalian cells.

In another preferred embodiment, the donor luciferase provided is *Renilla* luciferase. In yet another preferred embodiment, the acceptor fluorophore provided is an *Aequorea* green fluorescent protein.

In a particularly preferred embodiment, the detection of acceptor fluorophore fluorescence is performed using spectrofluorometry.

### DESCRIPTION

The present invention includes a method for determining whether a first protein interacts with a second protein in a living cell using luminescent resonance energy transfer (LRET). Luminescence resonance energy transfer results from the transfer of excited state energy from a donor luciferase to an acceptor fluorophore. In order for LRET to occur, there must be an overlap between the emission spectrum of the donor luciferase and the excitation spectrum of the acceptor fluorophore.

The efficiency of luminescence resonance energy transfer is dependent on the distance separating the donor luciferase and the acceptor fluorophore, among other variables.

Generally, significant energy transfers occur only where the donor luciferase and acceptor fluorophore are less than about 80 Å of each other. This short distance is considerably less than the distance needed between for optical resolution between two entities using conventional microscopy. Therefore, detecting luminescence resonance energy transfer between a donor luciferase and an acceptor fluorophore indicates that the donor luciferase and acceptor fluorophore have come within the distance needed for LRET to occur, that is less than about 80 Å of each other.

The present invention utilizes luminescence resonance energy transfer to determine whether an interaction takes place between a first protein and a second protein in a living cell. This is accomplished by complexing a first protein to the donor luciferase and complexing the second protein to the acceptor fluorophore and placing the complexed first protein and the complexed second protein in the cell under conditions suitable for an interaction between the first protein and the second protein to take place. If the first protein interacts with the second protein, the donor luciferase will come close enough to the acceptor fluorophore for luminescence resonance energy transfer to take place and the acceptor fluorophore will fluoresce. Detection of fluorescence from the acceptor fluorophore will, thereby, indicate that the first protein has interacted with the second protein.

Advantageously, this method allows for the detection of interaction between the first protein and the second protein even though the interaction cannot be detected by optical methods such as conventional microscopy.

There are several advantages of using luminescent resonance energy transfer to detect the interaction between two proteins according to the present invention. First, the specific labeling of the proteins in living cells can be achieved through genetic engineering methods where the introduction of fluorescent dyes into living cells is very difficult. Further, fluorescent dyes photobleach quickly while light emission of a luciferase such as *Renilla* luciferase originates from an enzymatic reaction that is relatively stable if substrate and oxygen are supplemented.

As used in this disclosure, "complexing a first protein to the donor luciferase" refers to joining the donor luciferase to the first protein in a manner that the donor luciferase and the first protein stay in essentially the same proximity to one another during interaction between the first protein and the second protein. Similarly, "complexing a second protein to

the acceptor fluorophore" refers to joining the acceptor fluorophore to the second protein in a manner that the acceptor fluorophore and the second protein stay in essentially the same proximity to one another during interaction between the first protein and the second protein. Such complexing can be done, for example, by genetically engineering the cell to produce a fusion protein containing the donor luciferase and first protein, and the acceptor fluorophore and the second protein.

In a preferred embodiment, the present invention uses *Renilla* luciferase as the donor luciferase and "humanized" *Aequorea* green fluorescent protein ('humanized' GFP) as the acceptor fluorophore. *Renilla* luciferase is a 34 kDa enzyme purified from *Renilla reniformis*. The enzyme catalyzes the oxidative decarboxylation of coelenterazine in the presence of oxygen to produce blue light with an emission wavelength maximum of 471 nm. *Renilla* luciferase was used as the donor luciferase because it requires an exogenous substrate rather than exogenous light for excitation. This, advantageously, eliminates background noise from an exogenous light source and from autofluorescence, and allows easy and accurate quantitative determination of light production.

'Humanized' GFP is a 27 kDa protein fluorophore that has an excitation maximum at 480 nm. It has a single amino acid difference from wild-type *Aequorea* green fluorescent protein. 'Humanized' GFP was chosen as the acceptor fluorophore because its excitation spectrum overlaps with the emission spectra of *Renilla* luciferase. Additionally, emissions from 'humanized' GFP can be visualized in living cells. Further, 'humanized' GFP is expressed well in the mammalian cells transfected with 'humanized' GFP cDNA that were used to demonstrate this method.

The method for determining whether a first protein interacts with a second protein according to the present invention was demonstrated as follows. In summary, insulin-like growth factor binding protein 6 (IGFBP 6) and insulin-like growth factor II (IGF-II) were selected as the first protein and second protein. IGFBP 6 is a protein known to have a marked binding affinity for IGF-II.

The *Renilla* luciferase cDNA was fused to IGFBP 6 cDNA and 'humanized' GFP cDNA was fused to IGF-II cDNA. Living cells were transfected with the fused cDNAs and the fusion proteins were expressed. Cell extracts were produced and mixed. The substrate for the *Renilla* luciferase moiety of the fused *Renilla* luciferase-IGFBP 6 protein

was added. Finally, fluorescence from the 'humanized' GFP moiety of the fused 'humanized' GFP-IGF-II protein was detected. Demonstration one method according to the present invention will now be described in greater detail.

**A) The Cloning of Fused IGFBP-6 cDNA to *Renilla* Luciferase cDNA; Fused IGF-II cDNA to 'humanized' GFP cDNA; and Fused Insulin cDNA to 'humanized' GFP cDNA:**

First, three fused cDNAs were produced: 1) fused IGFBP-6 cDNA and *Renilla* luciferase cDNA; 2) fused IGF-II cDNA and 'humanized' GFP cDNA; and 3) fused insulin cDNA and 'humanized' GFP cDNA. IGFBP-6 cDNA, SEQ ID NO:1, GenBank accession number M69054, encoded IGFBP-6, SEQ ID NO:2, which was used as the first protein.

*Renilla* luciferase cDNA, SEQ ID NO:3, GenBank accession number M63501, encoded *Renilla* luciferase, SEQ ID NO:4, which was used as the donor luciferase. IGF-II cDNA, SEQ ID NO:5, encoded IGF-II, SEQ ID NO:6, which was used as the second protein.

'Humanized' GFP cDNA, SEQ ID NO:7, GenBank accession number U50963, encoded 'humanized' GFP, SEQ ID NO:8, which was used as the acceptor fluorophore. Insulin cDNA, SEQ ID NO:9, accession number AH002844, encoded insulin, SEQ ID NO:10.

Insulin, fused to 'humanized' GFP, was used as a control protein because insulin is homologous to IGF-II, but it does not bind to IGFBP-6. The IGFBP-6 cDNA, SEQ ID NO:1, IGF-II cDNA, SEQ ID NO:5, and insulin cDNA, SEQ ID NO:9, were modified using PCR as follows.

First, the cDNA of prepro-IGF-II carried on an EcoRI fragment was cloned into pBluescript KS (+) II vector. The insert was sequenced using T7 and T3 primers and confirmed to contain the known cDNA sequence of prepro-IGF-II. The 5' end of the IGF-II precursor was connected to the T7 promoter in the pBluescript KS (+) II vector. An IGF-II 3' primer was designed to generate a Notice of Allowance restriction site, to remove the D and E domains of prepro-IGF-II, and to maintain the Notice of Allowance fragment of the 'humanized' GFP in frame with the open reading frame of IGF-II.

Next, the IGF-II fragment was amplified with PCR using the T7 promoter primer and the IGF-II 3' primer. The PCR-amplified IGF-II fragment was digested by EcoRI and Not I and cloned into pCDNA3.1 (+) vector (Invitrogen, Carlsbad, CA, US) producing pCDNA-IGF-II. Then, the Notice of Allowance fragment of the 'humanized' GFP was inserted into the Not I site of pCDNA-IGF-II producing pC-IGF-II-GFP.

The cDNA for precursor of insulin, which contained a signal peptide the B, C and A domains, was modified in a manner corresponding to the IGF-II fragment, above. The 'humanized' GFP cDNA was then linked to the 3' end of the modified insulin cDNA to produce pC-INS-GFP.

5                   Finally, IGFBP 6 cDNA was amplified by PCR from a plasmid named Rat-tagged human IGFBP6. The stop codon of IGFBP 6 was removed and the open reading frame of IGFBP 6 was in frame with *Renilla* luciferase cDNA from pCEP4-RUC (Mayerhofer R, Langridge WHR, Cormier MG and Szalay AA. *Expression of recombinant Renilla luciferase in transgenic plants results in high levels of light emission*. The Plant  
10   Journal 1995;7;1031-8). The linking of the *Renilla* luciferase cDNA to the 3' end of modified IGFBP 6 cDNA produced pC-IGFBP 6-RUC.

The sequences of the insert DNA fragments from all the constructs were verified by DNA sequencing analysis. Qiagen Maxi Plasmid Kit (Qiagen, Inc., Valencia, CA) was used for the purification of plasmid DNA.

15   **B) Transient Transfection of Mammalian Cells With pC-IGF-II-GFP, pC-INS-GFP and pC-IGFBP 6-RUC Using the Calcium Phosphate Precipitation Method:**

Next, mammalian cells were transfected with the cloned fusion DNAs. First, COS-7 cells (African green monkey kidney cell, American Type Culture Collection CRL 1651) were grown at 37 C in Dulbecco's Modified Eagle Medium (DMEM) with  
20   L-Glutamine supplemented with 10% fetal bovine serum and antibiotic antimycotic solution containing a final concentration of penicillin 100 unit/ml, streptomycin 100 mg/ml and amphotericin B 250 ng/ml (Sigma-Aldrich Co., St. Louis, MO, US) in 5% CO<sub>2</sub>. Groups of 1x10<sup>6</sup> of these cells were plated the day before transfection and were approximately 50% to 60% confluent at the time of transfection.

25                   Forty mg of each plasmid fusion DNA were precipitated and resuspended into Dulbecco's Phosphate Buffered Saline Solution and the plasmid fusion DNAs was introduced into mammalian cells using the standard calcium phosphate precipitation method. Transfection efficiency was estimated by fluorescence microscopy after 24 hours. The number of green fluorescent cells per plate were comparable in plates of pC-IGF-II-GFP  
30   DNA transfected cells, pC-INS-GFP DNA transfected cells and cells transfected with a plasmid DNA containing GFP only, which was used as a positive control.



**C) Confirmation of Expression of Fusion Proteins:**

Twenty-four hours after DNA transfection using DNA calcium phosphate precipitation method, individual plasmid DNA transfected COS-7 cells were visualized using fluorescence microscopy by detection of GFP fluorescence. pC-IGF-II-GFP and pC-INS-GFP transfected cells showed similar fluorescence patterns typical of secretory protein translocated through ER to Golgi. The pC-IGFBP 6-RUC transfected cells did not fluoresce. However, the pC-IGFBP 6-RUC transfected cells did show luminescence using a low light imaging system after the addition of coelenterazine.

Further, the presence of fusion proteins IGF-II-GFP and IGFBP 6-RUC, having the expected molecular weights of about 36 kDa and 56 kDa, respectively, were detected using immunoblot analysis. This confirmed the presence of both fusion proteins in the transiently transfected cells.

**D) Detection of Protein Interactions by Spectrofluorometry:**

Having confirmed the presence of the expected fusion proteins IGF-II-GFP and IGFBP 6-RUC, and the function of the donor luciferase and acceptor fluorophore, cell extracts from these transiently transfected cells were used to carry out a protein binding assay based on energy transfer between the *Renilla* luciferase and 'humanized' GFP moieties of the fusion proteins. Forty-eight hours after calcium transfection, the COS cells were washed twice with PBS and harvested using a cell scraper in luciferase assay buffer containing 0.5 M NaCl, 1 mM EDTA and 0.1 M potassium phosphate at a pH 7.5. The harvested cells were sonicated 3 times for 10 seconds with an interval of 10 seconds using a Fisher Model 550 Sonic Dismembrator (Fisher Scientific, Pittsburgh, PA, US) to produce cell extracts.

Next, the cell extracts containing IGF-II-GFP and IGFBP 6-RUC were mixed and 0.1  $\mu$ g of coelenterazine was immediately added. Spectrofluorometry was performed using a SPEX FluoroMax® (Instruments S.A., Inc., Edison, NJ). The spectrum showed a single emission peak at 471 nm, which corresponds to the known emission of *Renilla* luciferase.

Following the first spectrofluorometry, the mixtures were kept at room temperature for 30 minutes and the spectra were traced again after fresh coelenterazine was added. The trace at 30 minutes showed two peaks with emission maximum at 471 nm and 503 nm. The spectrofluorometry of the cell extracts was carried out at a longer time, but the

spectral pattern did not change over time.

Control cell extract mixtures from cells transfected with pC-INS-GFP and pC-IGFBP 6-RUC were made similarly and their spectra traced. The traces showed only one peak at 471 nm, which corresponds to the emission peak of *Renilla* luciferase. The spectral pattern did not change over time.

Therefore, these data demonstrated that IGFBP 6 and IGF-II interacted but that insulin and IGFBP 6 did not interact.

In addition to the above disclosed examples, protein-protein interactions were also detected by the detection of LRET using corresponding methods in *E. coli* cells and mammalian cells which were co-transformed.

Although the present invention has been discussed in considerable detail with reference to certain preferred embodiments, other embodiments are possible. For example, the interaction between molecules other than proteins could be studied by corresponding methods. Such other molecules could be provided to the living cell by diffusion, infusion, and incorporation or by other means. Further, fusion proteins produced from genetically engineered living cells could have post translational changes, such as the addition of sugar moieties, before their interactions are studied. Also, living cells can be visualized using these methods by spectrofluorometry by low light image analysis in cells, colonies and tissues. Additionally, high through put screening of colonies can be accomplished using the present methods combined with cell sorting and low light video analysis of micro titre dishes or multiple array detection. Therefore, the spirit and scope of the appended claims should not be limited to the description of preferred embodiments contained herein.

**WHAT IS CLAIMED IS:**

1. A method for determining whether a first protein interacts with a second protein within a living cell, the method comprising:

a) providing the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore within the cell;

b) placing the complexed first protein and the complexed second protein in proximity to each other within the cell; and

c) detecting any fluorescence from the acceptor fluorophore;

where the donor luciferase is capable of luminescence resonance energy transfer to the acceptor fluorophore when the first protein is in proximity to the second protein; and

where fluorescence of the acceptor fluorophore resulting from luminescence resonance energy transfer from the donor luciferase indicates that the first protein has interacted with the second protein.

2. The method of claim 1, where providing the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore comprises genetically engineering DNA and transferring the genetically engineered DNA to the living cell causing the cell to produce the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore.

3. The method of claim 1, where the cell provided with the first protein complexed to a donor luciferase is a mammalian cell.

4. The method of claim 1, where the cell provided with the second protein complexed to a acceptor fluorophore is a mammalian cell.

5. The method of claim 1, where the donor luciferase provided is *Renilla* luciferase.

6. The method of claim 1, where the acceptor fluorophore provided is a green fluorescent protein.

7. The method of claim 1, where the acceptor fluorophore provided is an *Aequorea* green fluorescent protein.

8. The method of claim 1, where detecting any fluorescence from the donor luciferase is performed using spectrofluorometry.

9. A method for determining whether a first molecule interacts with a second molecule within a living cell, the method comprising:

a) providing the first molecule complexed to a donor luciferase and the second molecule complexed to an acceptor fluorophore within the cell;

b) placing the complexed first molecule and the complexed second molecule in proximity to each other within the cell; and

5 c) detecting any fluorescence from the acceptor fluorophore;

where the donor luciferase is capable of luminescence resonance energy transfer to the acceptor fluorophore when the first molecule is in proximity to the second molecule; and

where fluorescence of the acceptor fluorophore resulting from luminescence resonance energy transfer from the donor luciferase indicates that the first molecule has interacted with  
10 the second molecule.

10. The method of claim 9, where the first molecule is a first protein and where the second molecule is a second protein; and

where providing the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore comprises genetically engineering DNA and  
15 transferring the genetically engineered DNA to the living cell causing the cell to produce the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore.

11. The method of claim 10, where the cell provided with the first protein complexed to a donor luciferase is a mammalian cell.

20 12. The method of claim 10, where the cell provided with the second protein complexed to a acceptor fluorophore is a mammalian cell.

13. The method of claim 9, where the donor luciferase provided is *Renilla* luciferase.

14. The method of claim 9, where the acceptor fluorophore provided is a green fluorescent protein.

25 15. The method of claim 9, where the acceptor fluorophore provided is a *Aequorea* green fluorescent protein.

16. The method of claim 9, where detecting any fluorescence from the donor luciferase is performed using spectrofluorometry.

17. A method for determining whether a first protein interacts with a second protein,  
30 the method comprising:

a) providing the first protein complexed to a donor luciferase and the second protein

complexed to an acceptor fluorophore;

b) placing the complexed first protein and the complexed second protein in proximity to each other; and

c) detecting any fluorescence from the acceptor fluorophore;

5 where the donor luciferase is capable of luminescence resonance energy transfer to the acceptor fluorophore when the first protein is in proximity to the second protein; and

where fluorescence of the acceptor fluorophore resulting from luminescence resonance energy transfer from the donor luciferase indicates that the first protein has interacted with the second protein.

10 18. The method of claim 17, where providing the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore comprises genetically engineering DNA and transferring the genetically engineered DNA to a living cell causing the cell to produce the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore.

15 19. The method of claim 18, where the cell provided with the first protein complexed to a donor luciferase is a mammalian cell.

20. The method of claim 18, where the cell provided with the second protein complexed to a acceptor fluorophore is a mammalian cell.

20 21. The method of claim 17, where the donor luciferase provided is *Renilla* luciferase.

22. The method of claim 17, where the acceptor fluorophore provided is a green fluorescent protein.

23. The method of claim 17, where the acceptor fluorophore provided is an *Aequorea* green fluorescent protein.

25 24. The method of claim 17, where detecting any fluorescence from the donor luciferase is performed using spectrofluorometry.

25. A method for determining whether a first molecule interacts with a second molecule, the method comprising:

30 a) providing the first molecule complexed to a donor luciferase and the second molecule complexed to an acceptor fluorophore;

b) placing the complexed first molecule and the complexed second molecule in

proximity to each other; and

c) detecting any fluorescence from the acceptor fluorophore;

where the donor luciferase is capable of luminescence resonance energy transfer to the acceptor fluorophore when the first molecule is in proximity to the second molecule; and

5        where fluorescence of the acceptor fluorophore resulting from luminescence resonance energy transfer from the donor luciferase indicates that the first molecule has interacted with the second molecule.

26. The method of claim 25, where the first molecule is a first protein and where the second molecule is a second protein; and

10        where providing the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore comprises genetically engineering DNA and transferring the genetically engineered DNA to a living cell causing the cell to produce the first protein complexed to a donor luciferase and the second protein complexed to an acceptor fluorophore.

15        27. The method of claim 26, where the cell provided with the first protein complexed to a donor luciferase is a mammalian cell.

28. The method of claim 26, where the cell provided with the second protein complexed to a acceptor fluorophore is a mammalian cell.

20        29. The method of claim 25, where the donor luciferase provided is *Renilla* luciferase.

30. The method of claim 25, where the acceptor fluorophore provided is a green fluorescent protein.

31. The method of claim 25, where the acceptor fluorophore provided is a *Aequorea* green fluorescent protein.

25        32. The method of claim 25, where detecting any fluorescence from the donor luciferase is performed using spectrofluorometry.

## SEQUENCE LISTING

&lt;110&gt; Szalay, Aladar A.

Wang, Yubao

Wang-Pruski, Gefu

Loma Linda University

&lt;120&gt; Method for Studying Protein Interactions in Vivo

&lt;130&gt; 11785-3PCT

&lt;140&gt; filed herewith

&lt;141&gt; 1999-09-02

&lt;150&gt; 60/135,835

&lt;151&gt; 1999-05-24

&lt;150&gt; 60/099,068

&lt;151&gt; 1998-09-03

&lt;160&gt; 10

&lt;170&gt; PatentIn Ver. 2.0

&lt;210&gt; 1

&lt;211&gt; 918

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (3)..(719)

&lt;400&gt; 1

tg tgc ccc cac agg ctg ctg cca ccg ctg ctg ctg cta gct ctg 47

Cys Pro His Arg Leu Leu Pro Pro Leu Leu Leu Leu Leu Ala Leu

1

5

10

15

ctg ctc gct gcc agc cca gga ggc gcc ttg gcg cgg tgc cca ggc tgc 95

Leu Leu Ala Ala Ser Pro Gly Gly Ala Leu Ala Arg Cys Pro Gly Cys

20

25

30

ggg caa ggg gtg cag gcg ggt tgt cca ggg ggc tgc gtg gag gag gag 143

Gly Gln Gly Val Gln Ala Gly Cys Pro Gly Gly Cys Val Glu Glu Glu

35

40

45

gat ggg ggg tcg cca gcc gag ggc tgc gcg gaa gct gag ggc tgt ctc 191

Asp Gly Gly Ser Pro Ala Glu Gly Cys Ala Glu Ala Glu Gly Cys Leu

50	55	60	
agg agg gag ggg cag gag tgc ggg gtc tac acc cct aac tgc gcc cca			239
Arg Arg Glu Gly Gln Glu Cys Gly Val Tyr Thr Pro Asn Cys Ala Pro			
65	70	75	
gga ctg cag tgc cat ccg ccc aag gac gac gag gcg cct ttg cgg gcg			287
Gly Leu Gln Cys His Pro Pro Lys Asp Asp Glu Ala Pro Leu Arg Ala			
80	85	90	95
ctg ctg ctc ggc cga ggc cgc tgc ctt ccg gcc cgc gcg cct gct gtt			335
Leu Leu Leu Gly Arg Gly Arg Cys Leu Pro Ala Arg Ala Pro Ala Val			
100	105	110	
gca gag gag aat cct aag gag agt aaa ccc caa gca ggc act gcc cgc			383
Ala Glu Glu Asn Pro Lys Glu Ser Lys Pro Gln Ala Gly Thr Ala Arg			
115	120	125	
cca cag gat gtg aac cgc aga gac caa cag agg aat cca ggc acc tct			431
Pro Gln Asp Val Asn Arg Arg Asp Gln Gln Arg Asn Pro Gly Thr Ser			
130	135	140	
acc acg ccc tcc cag ccc aat tct gcg ggt gtc caa gac act gag atg			479
Thr Thr Pro Ser Gln Pro Asn Ser Ala Gly Val Gln Asp Thr Glu Met			
145	150	155	
ggc cca tgc cgt aga cat ctg gac tca gtg ctg cag caa ctc cag act			527
Gly Pro Cys Arg Arg His Leu Asp Ser Val Leu Gln Gln Leu Gln Thr			
160	165	170	175
gag gtc tac cga ggg gct caa aca ctc tac gtg ccc aat tgt gac cat			575
Glu Val Tyr Arg Gly Ala Gln Thr Leu Tyr Val Pro Asn Cys Asp His			
180	185	190	
cga ggc ttc tac cgg aag cgg cag tgc cgc tcc tcc cag ggg cag cgc			623
Arg Gly Phe Tyr Arg Lys Arg Gln Cys Arg Ser Ser Gln Gly Gln Arg			
195	200	205	
cga ggt ccc tgc tgg tgt gtg gat cgg atg ggc aag tcc ctg cca ggg			671
Arg Gly Pro Cys Trp Cys Val Asp Arg Met Gly Lys Ser Leu Pro Gly			
210	215	220	
tct cca gat ggc aat gga agc tcc tcc tgc ccc act ggg agt agc ggc			719
Ser Pro Asp Gly Asn Gly Ser Ser Ser Cys Pro Thr Gly Ser Ser Gly			
225	230	235	
taaagctggg ggatagaggg gctgcagggc cactggaagg aacatggagc tgtcatcact			779



caacaaaaaa ccgagggcct caatccacct tcagggcccg ccccatgggc ccctcaccgc 839  
 tgggttgaaa gagtgttggt gttggctggg gtgtcaataa agctgtgctt ggggtcgctg 899  
 aaaaaaaaaa aaaaaaaaaa 918

<210> 2

<211> 239

<212> PRT

<213> Homo sapiens

<400> 2

Cys Pro His Arg Leu Leu Pro Pro Leu Leu Leu Leu Leu Ala Leu Leu  
 1 5 10 15

Leu Ala Ala Ser Pro Gly Gly Ala Leu Ala Arg Cys Pro Gly Cys Gly  
 20 25 30

Gln Gly Val Gln Ala Gly Cys Pro Gly Gly Cys Val Glu Glu Glu Asp  
 35 40 45

Gly Gly Ser Pro Ala Glu Gly Cys Ala Glu Ala Glu Gly Cys Leu Arg  
 50 55 60

Arg Glu Gly Gln Glu Cys Gly Val Tyr Thr Pro Asn Cys Ala Pro Gly  
 65 70 75 80

Leu Gln Cys His Pro Pro Lys Asp Asp Glu Ala Pro Leu Arg Ala Leu  
 85 90 95

Leu Leu Gly Arg Gly Arg Cys Leu Pro Ala Arg Ala Pro Ala Val Ala  
 100 105 110

Glu Glu Asn Pro Lys Glu Ser Lys Pro Gln Ala Gly Thr Ala Arg Pro  
 115 120 125

Gln Asp Val Asn Arg Arg Asp Gln Gln Arg Asn Pro Gly Thr Ser Thr  
 130 135 140

Thr Pro Ser Gln Pro Asn Ser Ala Gly Val Gln Asp Thr Glu Met Gly  
 145 150 155 160

Pro Cys Arg Arg His Leu Asp Ser Val Leu Gln Gln Leu Gln Thr Glu  
 165 170 175

Val Tyr Arg Gly Ala Gln Thr Leu Tyr Val Pro Asn Cys Asp His Arg  
 180 185 190

Gly Phe Tyr Arg Lys Arg Gln Cys Arg Ser Ser Gln Gly Gln Arg Arg  
 195 200 205

Gly Pro Cys Trp Cys Val Asp Arg Met Gly Lys Ser Leu Pro Gly Ser  
 210 215 220

Pro Asp Gly Asn Gly Ser Ser Ser Cys Pro Thr Gly Ser Ser Gly  
 225 230 235

<210> 3

<211> 1196

<212> DNA

<213> *Renilla reniformis*

<220>

<221> CDS

<222> (10)..(945)

<400> 3

agcttaaag atg act tcg aaa gtt tat gat cca gaa caa agg aaa cgg atg 51  
 Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met  
 1 5 10

ata act ggt ccg cag tgg tgg gcc aga tgt aaa caa atg aat gtt ctt 99  
 Ile Thr Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu  
 15 20 25 30

gat tca ttt att aat tat tat gat tca gaa aaa cat gca gaa aat gct 147  
 Asp Ser Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala  
 35 40 45

gtt att ttt tta cat ggt aac gcg gcc tct tct tat tta tgg cga cat 195  
 Val Ile Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His  
 50 55 60

gtt gtg cca cat att gag cca gta gcg cgg tgt att ata cca gat ctt 243  
 Val Val Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu  
 65 70 75

att ggt atg ggc aaa tca ggc aaa tct ggt aat ggt tct tat agg tta 291  
 Ile Gly Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu  
 80 85 90

ctt gat cat tac aaa tat ctt act gca tgg ttt gaa ctt ctt aat tta 339  
 Leu Asp His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu  
 95 100 105 110

```

cca aag aag atc att ttt gtc ggc cat gat tgg ggt gct tgt ttg gca 387
Pro Lys Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala
      115              120              125

ttt cat tat agc tat gag cat caa gat aag atc aaa gca ata gtt cac 435
Phe His Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val His
      130              135              140

gct gaa agt gta gta gat gtg att gaa tca tgg gat gaa tgg cct gat 483
Ala Glu Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp
      145              150              155

att gaa gaa gat att gcg ttg atc aaa tct gaa gaa gga gaa aaa atg 531
Ile Glu Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met
      160              165              170

gtt ttg gag aat aac ttc ttc gtg gaa acc atg ttg cca tca aaa atc 579
Val Leu Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile
      175              180              185              190

atg aga aag tta gaa cca gaa gaa ttt gca gca tat ctt gaa cca ttc 627
Met Arg Lys Leu Glu Pro Glu Glu Phe Ala Ala Tyr Leu Glu Pro Phe
      195              200              205

aaa gag aaa ggt gaa gtt cgt cgt cca aca tta tca tgg cct cgt gaa 675
Lys Glu Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu
      210              215              220

atc ccg tta gta aaa ggt ggt aaa cct gac gtt gta caa att gtt agg 723
Ile Pro Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg
      225              230              235

aat tat aat gct tat cta cgt gca agt gat gat tta cca aaa atg ttt 771
Asn Tyr Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe
      240              245              250

att gaa tcg gat cca gga ttc ttt tcc aat gct att gtt gaa ggc gcc 819
Ile Glu Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala
      255              260              265              270

aag aag ttt cct aat act gaa ttt gtc aaa gta aaa ggt ctt cat ttt 867
Lys Lys Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe
      275              280              285

tcg caa gaa gat gca cct gat gaa atg gga aaa tat atc aaa tcg ttc 915
Ser Gln Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe
      290              295              300

```

gtt gag cga gtt ctc aaa aat gaa caa taa ttactttggg tttttattta 965  
 Val Glu Arg Val Leu Lys Asn Glu Gln  
 305 310

catttttccc gggtttaata atataaatgt cattttcaac aattttatatt taactgaata 1025  
 tttcacaggg aacattcata tatgttgatt aatttagctc gaactttact ctgtcatatc 1085  
 attttggaat attacctctt tcaatgaaac ttataaaca gtgggttcaat taattaatat 1145  
 atattataat tacatttggt atgtaataaa ctcggtttta ttataaaaaa a 1196

<210> 4  
 <211> 311  
 <212> PRT  
 <213> Renilla reniformis

<400> 4  
 Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg Met Ile Thr  
 1 5 10 15  
 Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val Leu Asp Ser  
 20 25 30  
 Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn Ala Val Ile  
 35 40 45  
 Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg His Val Val  
 50 55 60  
 Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp Leu Ile Gly  
 65 70 75 80  
 Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg Leu Leu Asp  
 85 90 95  
 His Tyr Lys Tyr Leu Thr Ala Trp Phe Glu Leu Leu Asn Leu Pro Lys  
 100 105 110  
 Lys Ile Ile Phe Val Gly His Asp Trp Gly Ala Cys Leu Ala Phe His  
 115 120 125  
 Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val His Ala Glu  
 130 135 140  
 Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro Asp Ile Glu

145                                      150                                      155                                      160  
 Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys Met Val Leu  
    165                                      170                                      175  
 Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys Ile Met Arg  
    180                                      185                                      190  
 Lys Leu Glu Pro Glu Glu Phe Ala Ala Tyr Leu Glu Pro Phe Lys Glu  
    195                                      200                                      205  
 Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg Glu Ile Pro  
    210                                      215                                      220  
 Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val Arg Asn Tyr  
    225                                      230                                      235                                      240  
 Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met Phe Ile Glu  
    245                                      250                                      255  
 Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly Ala Lys Lys  
    260                                      265                                      270  
 Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His Phe Ser Gln  
    275                                      280                                      285  
 Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser Phe Val Glu  
    290                                      295                                      300  
 Arg Val Leu Lys Asn Glu Gln  
 305                                      310  
  
 <210> 5  
 <211> 543  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> CDS  
 <222> (1)..(543)  
  
 <400> 5  
 atg gga atc cca atg ggg aag tcg atg ctg gtg ctt ctc acc ttc ttg    48  
 Met Gly Ile Pro Met Gly Lys Ser Met Leu Val Leu Leu Thr Phe Leu  
      1                                      5                                      10                                      15  
 gcc ttc gcc tcg tgc tgc att gct gct tac cgc ccc agt gag acc ctg    96

Ala Phe Ala Ser Cys Cys Ile Ala Ala Tyr Arg Pro Ser Glu Thr Leu  
                   20                                  25                                  30

tgc ggc ggg gag ctg gtg gac acc ctc cag ttc gtc tgt ggg gac cgc 144  
 Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val Cys Gly Asp Arg  
                   35                                  40                                  45

ggc ttc tac ttc agc agg ccc gca agc cgt gtg agc cgt cgc agc cgt 192  
 Gly Phe Tyr Phe Ser Arg Pro Ala Ser Arg Val Ser Arg Arg Ser Arg  
                   50                                  55                                  60

ggc atc gtt gag gag tgc tgt ttc cgc agc tgt gac ctg gcc ctc ctg 240  
 Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala Leu Leu  
                   65                                  70                                  75                                  80

gag acg tac tgt gct acc ccc gcc aag tcc gag agg gac gtg tcg acc 288  
 Glu Thr Tyr Cys Ala Thr Pro Ala Lys Ser Glu Arg Asp Val Ser Thr  
                                   85                                  90                                  95

cct ccg acc gtg ctt ccg gac aac ttc ccc aga tac ccc gtg ggc aag 336  
 Pro Pro Thr Val Leu Pro Asp Asn Phe Pro Arg Tyr Pro Val Gly Lys  
                                   100                                  105                                  110

ttc ttc caa tat gac acc tgg aag cag tcc acc cag cgc ctg cgc agg 384  
 Phe Phe Gln Tyr Asp Thr Trp Lys Gln Ser Thr Gln Arg Leu Arg Arg  
                   115                                  120                                  125

ggc ctg cct gcc ctc ctg cgt gcc cgc cgg ggt cac gtg ctc gcc aag 432  
 Gly Leu Pro Ala Leu Leu Arg Ala Arg Arg Gly His Val Leu Ala Lys  
                   130                                  135                                  140

gag ctc gag gcg ttc agg gag gcc aaa cgt cac cgt ccc ctg att gct 480  
 Glu Leu Glu Ala Phe Arg Glu Ala Lys Arg His Arg Pro Leu Ile Ala  
 145                                  150                                  155                                  160

cta ccc acc caa gac ccc gcc cac ggg ggc gcc ccc cca gag atg gcc 528  
 Leu Pro Thr Gln Asp Pro Ala His Gly Gly Ala Pro Pro Glu Met Ala  
                                   165                                  170                                  175

agc aat cgg aag tga 543  
 Ser Asn Arg Lys  
                   180

&lt;210&gt; 6

&lt;211&gt; 180

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 6

Met Gly Ile Pro Met Gly Lys Ser Met Leu Val Leu Leu Thr Phe Leu  
 1 5 10 15

Ala Phe Ala Ser Cys Cys Ile Ala Ala Tyr Arg Pro Ser Glu Thr Leu  
 20 25 30

Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val Cys Gly Asp Arg  
 35 40 45

Gly Phe Tyr Phe Ser Arg Pro Ala Ser Arg Val Ser Arg Arg Ser Arg  
 50 55 60

Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala Leu Leu  
 65 70 75 80

Glu Thr Tyr Cys Ala Thr Pro Ala Lys Ser Glu Arg Asp Val Ser Thr  
 85 90 95

Pro Pro Thr Val Leu Pro Asp Asn Phe Pro Arg Tyr Pro Val Gly Lys  
 100 105 110

Phe Phe Gln Tyr Asp Thr Trp Lys Gln Ser Thr Gln Arg Leu Arg Arg  
 115 120 125

Gly Leu Pro Ala Leu Leu Arg Ala Arg Arg Gly His Val Leu Ala Lys  
 130 135 140

Glu Leu Glu Ala Phe Arg Glu Ala Lys Arg His Arg Pro Leu Ile Ala  
 145 150 155 160

Leu Pro Thr Gln Asp Pro Ala His Gly Gly Ala Pro Pro Glu Met Ala  
 165 170 175

Ser Asn Arg Lys  
 180

&lt;210&gt; 7

&lt;211&gt; 717

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(717)

&lt;220&gt;

<223> Description of Artificial Sequence: humanized  
green fluorescent protein cDNA

&lt;400&gt; 7

```

atg agc aag ggc gag gaa ctg ttc act ggc gtg gtc cca att ctc gtg   48
Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
  1             5             10             15

gaa ctg gat ggc gat gtg aat ggg cac aaa ttt tct gtc agc gga gag   96
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
      20             25             30

ggg gaa ggt gat gcc aca tac gga aag ctc acc ctg aaa ttc atc tgc  144
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
      35             40             45

acc act gga aag ctc cct gtg cca tgg cca aca ctg gtc act acc ttc  192
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
      50             55             60

tct tat ggc gtg cag tgc ttt tcc aga tac cca gac cat atg aag cag  240
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
      65             70             75             80

cat gac ttt ttc aag agc gcc atg ccc gag ggc tat gtg cag gag aga  288
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
      85             90             95

acc atc ttt ttc aaa gat gac ggg aac tac aag acc cgc gct gaa gtc  336
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
      100            105            110

aag ttc gaa ggt gac acc ctg gtg aat aga atc gag ctg aag ggc att  384
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
      115            120            125

gac ttt aag gag gat gga aac att ctc ggc cac aag ctg gaa tac aac  432
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
      130            135            140

tat aac tcc cac aat gtg tac atc atg gcc gac aag caa aag aat ggc  480
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
      145            150            155            160

atc aag gtc aac ttc aag atc aga cac aac att gag gat gga tcc gtg  528
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
      165            170            175

```



cag ctg gcc gac cat tat caa cag aac act cca atc ggc gac ggc cct 576  
 Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro  
 180 185 190

gtg ctc ctc cca gac aac cat tac ctg tcc acc cag tct gcc ctg tct 624  
 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser  
 195 200 205

aaa gat ccc aac gaa aag aga gac cac atg gtc ctg ctg gag ttt gtg 672  
 Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val  
 210 215 220

acc gct gct ggg atc aca cat ggc atg gac gag ctg tac aag tga 717  
 Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys  
 225 230 235

<210> 8

<211> 238

<212> PRT

<213> Artificial Sequence

<400> 8

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val  
 1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu  
 20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys  
 35 40 45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe  
 50 55 60

Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln  
 65 70 75 80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg  
 85 90 95

Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val  
 100 105 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile  
 115 120 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn  
 130 135 140

Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly  
 145 150 155 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val  
 165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro  
 180 185 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser  
 195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val  
 210 215 220

Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys  
 225 230 235

<210> 9

<211> 333

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(333)

<400> 9

atg gcc ctg tgg atg cgc ctc ctg ccc ctg ctg gcg ctg ctg gcc ctc 48  
 Met Ala Leu Trp Met Arg Leu Leu Pro Leu Leu Ala Leu Leu Ala Leu  
 1 5 10 15

tgg gga cct gac cca gcc gca gcc ttt gtg aac caa cac ctg tgc ggc 96  
 Trp Gly Pro Asp Pro Ala Ala Ala Phe Val Asn Gln His Leu Cys Gly  
 20 25 30

tca cac ctg gtg gaa gct ctc tac cta gtg tgc ggg gaa cga ggc ttc 144  
 Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe  
 35 40 45

ttc tac aca ccc aag acc cgc cgg gag gca gag gac ctg cag gtg ggc 192  
 Phe Tyr Thr Pro Lys Thr Arg Arg Glu Ala Glu Asp Leu Gln Val Gly  
 50 55 60

cag gtg gag ctg ggc ggg ggc cct ggt gca ggc agc ctg cag ccc ttg 240  
 Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu  
 65 70 75 80

gcc ctg gag ggg tcc ctg cag aag cgt ggc att gtg gaa caa tgc tgt 288  
 Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu Gln Cys Cys  
 85 90 95

acc agc atc tgc tcc ctc tac cag ctg gag aac tac tgc aac tag 333  
 Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn  
 100 105 110

<210> 10

<211> 110

<212> PRT

<213> Homo sapiens

<400> 10

Met Ala Leu Trp Met Arg Leu Leu Pro Leu Leu Ala Leu Leu Ala Leu  
 1 5 10 15

Trp Gly Pro Asp Pro Ala Ala Ala Phe Val Asn Gln His Leu Cys Gly  
 20 25 30

Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe  
 35 40 45

Phe Tyr Thr Pro Lys Thr Arg Arg Glu Ala Glu Asp Leu Gln Val Gly  
 50 55 60

Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu  
 65 70 75 80

Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu Gln Cys Cys  
 85 90 95

Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn  
 100 105 110

## INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US99/20207

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

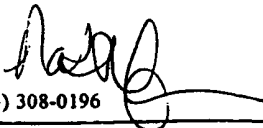
U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
NONEElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
WEST, CAS ONLINE**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y, E	US 5,976,796 A (SZALAY et al.) 02 November 1999, col. 1, line 60-col. 3, line 17.	1-32
Y	US 5,491,084 A (CHALFIE et al.) 13 February 1996, col. 1, lines 36-61.	1-32
Y	US 5,418,155 A (CORMIER et al.) 23 May 1995, col. 2, line 59-col. 10, line 9.	1-32
Y, P	US 5,891,646 A (BARAK et al.) 06 April 1999, col. 2, line 6-col. 4, line 47.	1-32
A	US 5,866,348 A (SCHEIRER) 02 February 1999, col. 2, line 15-col. 3, line 27.	1-32

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 06 NOVEMBER 1999	Date of mailing of the international search report <b>19 NOV 1999</b>
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer PENSEE T. DO  Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/20207

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4,604,364 A (KOSAK) 05 August 1986, col. 3, line 26-col. 8, line 14.	1-32
A	US 4,318,707 A (LITMAN et al.) 09 March 1982, col. 3, line 1-col. 7, line 14.	1-32
A	US 5,683,888 A (CAMPBELL) 04 November 1997, col. 2, line 10-col. 4, line 50.	1-32
A, P	US 5,811,238 A (STEMMER et al.) 22 September 1998, col. 4, line 11-col. 7, line 35.	1-32

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US99/20207

## A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

C12Q 1/66, 1/00; C12P 21/06; C07H 21/04; G01N 33/53, 33/567, 33/573, 33/58, 33/60, 33/531, 33/533, 21/00, 21/76, 21/75, 31/00

## A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

435/4, 7.1, 7.2, 7.21, 7.4, 7.91, 8, 68.1, 69.1, 70.1, 70.3, 127, 231, 232, 326, 331, 366, 375, 968, 972; 436/543, 546, 164, 172, 800, 805, 905; 536/23.5; 424/9.6

## B. FIELDS SEARCHED

Minimum documentation searched

Classification System: U.S.

435/4, 7.1, 7.2, 7.21, 7.4, 7.91, 8, 68.1, 69.1, 70.1, 70.3, 127, 231, 232, 326, 331, 366, 375, 968, 972; 436/543, 546, 164, 172, 800, 805, 905; 536/23.5; 424/9.6